

~~SECRET~~

NOFORN

Approved For Release 2002/01/08 : CIA-RDP83T00573R000100120011-9

Cross-Program

ADP

Issue:

Should funds be included in the FY 1981 NFIP budget request to implement a Community-wide Bibliographic and Document Storage and Retrieval System?

Reason for Issue:

- o Studies completed by the IHC during 1979 strongly support the conclusion that a requirement for this system exists in the Intelligence Community. Implementation could significantly reduce information handling costs by eliminating unnecessary duplication of effort in the future.

- o Because of its Community-wide application, this system should be funded on a NFIP-wide basis, and the initiative for inclusion in the FY 1981 budget should properly come from RMS rather than any of the NFIP departments and agencies affected.

Value of the Activity:

- o The overwhelming majority of ADP resources employed by the Intelligence Community are devoted to the collection, processing, transmission, cataloguing, and storing of intelligence information. Expenditures supporting these activities achieve maximum value only to the degree that analytical outputs are delivered to production organizations in a timely and usable manner.

- o To the extent that adequate indexes to existing information are not available to producers, the Intelligence Community's product will not take into account all of the pertinent information which has been collected and processed. A significant portion of the total collection and processing effort will thereby be wasted and the quality of analysis will be degraded.

- o The primary users of the proposed system will be the analytical components of State, DIA, NSA, Treasury, and the Military Services.

Objectives of the Alternatives:

The objective of the alternatives is to implement a Community Bibliographic and Document Storage and Retrieval System commencing in FY 1981. A secondary but important objective of Alternative 1

~~SECRET~~

NOFORN

Approved For Release 2002/01/08 : CIA-RDP83T00573R000100120011-9

Approved For Release 2002/01/08 : CIA-RDP83T00573R000100120011-9

is to provide facilities with which CIA can support Community systems which have been proposed, such as the 4C system for security clearances and systems to support the work of the Interagency Defector Committee and the Committee on Exchanges.

Description of Alternatives:

Alternative 1: Provide direct on-line access by the Community to CIA's RECON bibliographic index using COINS. Fiscal impact of the principal funding alternatives are as follows:

\$000	FY 81	FY 82	FY 83	FY 84	FY 85
1a	1,760	1,310	1,310	1,310	1,310
1b	1,611	686	686	686	686
1c	375	1,428	1,428	1,428	1,428
1d	375	720	720	720	720
Personnel:					
1a & 1b	12	15	15	15	15
1c & 1d	3	15	15	15	15

Alternative 2: Provide direct on-line access to CIA's RECON bibliographic index using intermediary systems. Fiscal impact:

	FY 81	FY 82	FY 83	FY 84	FY 85
\$000	950	400	400	400	400

Alternative 3: Include the requirement for this system as part of the CIA-DIA SAFE Project. Add resources to the FY 1982 SAFE budget in the CIAP as necessary.

Alternative 4: Defer all action on the CIA proposal pending further study.

Analysis

o In 1974 ASD(I) recommended that CIA's RECON bibliographic index be made accessible throughout the Community using COINS. To date, only a small portion of the CIA's automated index is available directly to the Community by duplication of portions of the data base on the NSA COINS host.

o The June 1979 Report of the Analyst Support Task Force, which was sponsored by the DCI's Intelligence Information Handling Committee (IHC), concluded that "a Community-wide, on-line, multi-source bibliographic index to Intelligence Community documents should be created and made available to the Intelligence Community's production analysts."

Approved For Release 2002/01/08 : CIA-RDP83T00573R000100120011-9

o While the Analyst Support Study was being conducted, the CIA submitted a proposal to the Chairman IHC to make its all-source bibliographic reference file (RECON) available to the Community. The CIA proposal included its newly developed ADSTAR automated microfilm document storage and retrieval system for use, together with RECON as a centralized all-source Community bibliographic index and document storage and retrieval system. A copy of the CIA proposal is at Tab A.

o Acting on the recommendation of the Analyst Support Task Group and in response to the CIA proposal, a study of the advisability of implementing all or part of the CIA proposal is being conducted by the IHC. Although the portion of the CIA system dealing with ADSTAR is not complete, a preliminary recommendation concerning RECON has been formulated and is contained in the preliminary report to the IHC, a copy of which is at Tab B. The report recommends that the Community proceed to implement a modification of the CIA proposal as part of an overall plan to create a Community Bibliographic Index and Document Storage and Retrieval System.

o The IHC study demonstrates that providing Community-wide access to the RECON data base would enhance information retrieval capabilities of Community analysts outside CIA. A comparison of RECON with other Community systems demonstrates that it possesses both greater breadth (it includes index records to over 50% of all of the classified intelligence information and reports generated by the Community), and greater depth (it covers a time span of 10 years) than any other existing bibliographic system. Its inclusion, along with more specialized bibliographic type systems such as SOLIS, NDS, and the IRISA and ASDIA files on DIAOLS, among the automated index systems accessible through COINS would provide the foundation upon which to build a comprehensive all-source Community Bibliographic Index and Document Storage and Retrieval System.

o Alternative 1: Direct On-Line Service. As stated in the CIA proposal: "The principal advantages of this arrangement include the significantly faster availability of the document citation to the analyst, plus the capability of the analyst to work directly with the data base." Other significant and important advantages would also accrue from this approach. By placing RECON on COINS, CIA analysts could be provided direct access to all COINS data bases. In addition, if a computer of sufficient capacity is provided (financing options 1(a) and 1(c)) other Community systems could be implemented without substantial additional costs being incurred for computer hardware, operation, or maintenance.

o Four funding options are included under Alternative 1 for inclusion in the FY 1981 CIAP and in the COINS decision unit in the CCP. The basis for each is discussed in attachment, Tab C.

- 1(a) Provides funds for acquisition of a new Amdahl V-5 (or equivalent) CPU and peripheral equipment together with software development funds required for implementation in FY 81.
- 1(b) Provides for implementation by reverting to the originally planned schedule for acquiring the CAMS back-up CPU and utilizing the present back-up 370/158 CPU to support RECON.
- 1(c) Provides minimum software development funds in FY 1981 and programs equipment acquisition as in 1(a) commencing in FY 1982.
- 1(d) Provides minimum software development funds in FY 1981 and utilizes the surplus CAMS 370/158 back-up CPU commencing in FY 82.

Option 1(a) provides for implementation at the earliest possible date with sufficient upgrade capacity to support other Community applications. Option 1(b) permits RECON to be implemented as soon as possible with the lowest life cycle cost by avoiding the necessity of acquiring new CPU for this purpose. It has the disadvantage of limiting the capacity of the system to support other Community applications and diverting a COINS network access system funded in FY 80 from other possible applications. Option 1(c) provides the same capabilities as 1(a) with minimum funding requirements in FY 81. It has the disadvantage of postponing the availability of the system until FY 82. Option 1(d) has the advantage of requiring the minimum level of funding in both FY 81 and the outyears, but has the disadvantage of postponing system availability until FY 82, limiting the flexibility and growth potential of the system, and requires the use of the present 370/158 CAMS back-up CPU which could be utilized for other purposes.

- o If either 1(c) or 1(d) is chosen, the decision to proceed could be reconsidered during the 1982 Program Review after the final report to the IHC by the working group has been considered, but prior to the expenditure of the budgeted funds.

- o Implementation of Alternative 1 will require a maximum of 1500 square feet of additional computer grade floor space at CIA Headquarters in or adjacent to an existing CIA computer center. It is believed this amount of space can be accommodated in or adjacent to existing CIA computer centers, within the site being prepared for the SAFE project, or within space coming available at CIA Headquarters due to new GSA leases. If it is not possible to secure space in or adjacent to existing centers, implementation may require up to 2,500 square feet of space.

Approved For Release 2002/01/08 : CIA-RDP83T00573R000100120011-9

o Alternative 2 has the advantage of making the RECON system available to the Community at lower cost, but on a limited basis. It provides no additional capabilities to CIA and provides no opportunities for CIA to support other Community systems or applications.

o What is proposed for implementation under either Alternatives 1 or 2 is limited to the existing RECON data base, rather than the expansion included in the CIA proposal. The analysis completed by the IHC is insufficient to justify the additional costs in money and personnel which would be required to expand the scope of the RECON and ADSTAR systems. Should further study and analysis provide such a justification, these features could be added later if additional resources are provided.

o Alternative 3 would defer any substantial expenditures to provide Community services in the bibliographic area but attempts to ensure that the objectives of the proposed system will ultimately be realized using SAFE. However, requiring and relying upon SAFE to provide these Community services could introduce an element of technical risk into the SAFE project which could cause a delay in the execution of the entire project with concomitant cost overruns.

o Alternative 4 would permit implementation to be done after more thorough planning and analysis of all aspects of the CIA proposal, including ADSTAR, has taken place. However, actual implementation may be thereby delayed sufficiently long to preclude implementation at all based upon the limits of the life expectancy of the system.

ODP-8-2184

25 JAN 1979

MEMORANDUM FOR: Chairman, DCI Intelligence Information
Handling Committee

FROM : Clifford D. May, Jr.
CIA Member, IHC

SUBJECT : Proposal for a Centralized Community
Bibliographic and Document Retrieval
System Operated by CIA

1. Proposal: This memorandum proposes that Intelligence Information Handling Committee study the feasibility and desirability of adopting CIA's RECON bibliographic index and ADSTAR micrographic document storage and retrieval system as a Centralized Intelligence Community Bibliographic and Document Retrieval System, managed and operated for the Community by CIA.

2. Background: a. The RECON subject file, from which the proposed Community data base would be derived, has several advantages over other computer-based document indexing systems currently used by NFIB agencies. Initiated in 1968, the RECON file is the largest and most comprehensive subject index to intelligence reports in the Community. As of September 1978 the file contained 3,000,000 index records. RECON offers access to virtually all substantive intelligence documents originated (given general distribution) by the CIA, DoD, DIA, Air Force, Army, Navy, NSA, State, and NPIC, and some documents from other government agencies of the United States [REDACTED]. The data base contains both raw and finished intelligence reports, includes both collateral intelligence and Sensitive Compartmented Information (SCI), and the area coverage is worldwide. Subjects indexed include government, politics, society, culture, science and technology, transportation, communications, business, commerce, industry, finance, commodities (both strategic and non-strategic), products (civilian and military), resources (including labor and military manpower), and the armed forces. In brief, no area of interest to intelligence is overlooked. Open literature, non-CIA cables, and [REDACTED] reporting are included on a selective basis.

STATINTL

STATSPEC

b. The full RECON data base is stored in machine-readable form and is searchable by computer via any one or a combination of the elements used to describe each document. These include the bibliographic description (title, issuing agency, post or origin, date, report number, security classification and dissemination restrictions); area codes (China and the Soviet Union are subdivided to the province and oblast level, respectively); specific place names where appropriate; subject codes; and keywords. The 320 subject codes are standardized broad subdivisions, more than one of which can be assigned to any single document by the indexers in CIA's Office of Central Reference (OCR). The keywords are non-standardized terms added by the indexer based on review of the title and document text; these individual keywords supplement the broader subject codes and thus refine the retrievability of each individual document. The flexibility of such an indexing system allows it to easily accommodate new subject indexing requirements.

c. RECON has an historical depth of 10 years and is the most up-to-date general purpose subject index to intelligence documents available. Approximately 85-90 percent of incoming documents are available for computer search of the index records within eight days after receipt, and by July 1979 this figure will be reduced to three days. Portions of the RECON data base are now available to the Community via COINS, and the total data base itself has been queried on a limited basis by OCR analysts for all NFIB agencies continually since its development. When CIA's earlier bibliographic retrieval system, known as "Intellofax," was in operation, then non-CIA use of the CIA index to intelligence reports was about 45 percent of total queries. With the initiation of the AEGIS/RECON system in 1967-68, however, CIA management placed severe limits on other agency access to these bibliographic records because of substantial reductions imposed on CIA resources. Even under this restriction, however, non-CIA use of the data base has crept upward, and during the first half of CY-1978 the entire data base was queried over 800 times by non-CIA NFIB agencies (approximately 26% of total queries during this period). During the same period, the finished intelligence portion of the RECON data base, which is part of the COINS system, was queried via COINS by non-CIA NFIB agencies over 1,200 times.

d. Bibliographic services must be supplemented by document retrieval capabilities. To ensure speedy and efficient retrieval, CIA is building an Automated Document Storage and Retrieval (ADSTAR) System, which is scheduled to enter operation in November 1979. Designed to operate either in batch or online mode, ADSTAR will store documents on microfilm but digitize these images for transmission over broad-band communications links to remote display terminals and printers.

3. Community Options for Bibliographic Service:

a. Offline Service

(1) The least costly approach of providing RECON bibliographic records to the Community would simply entail offering increased service from the system in its present configuration to other NFIB members. Under this arrangement, a non-CIA analyst presents his research request in writing or over the phone to an OCR area reference analyst, who queries the RECON data base and then mails the printed listing of records to the original requester.

(2) The primary disadvantages of this system are the delays involved in having to mail the request and document listing. The existence of an intermediary (the OCR area reference analyst) between the end user of the data and the data base itself can also be a disadvantage, but not without some positive aspects. Among the disadvantages, the requester may have no way of knowing how large or small a document listing he will be getting until he receives it from the area reference analyst. Any revision of his query to make his request either more inclusive, more selective, or otherwise more appropriate for retrieving precisely what he needs can only be made after the query has been run and the complete document listing is received through the mail. On the positive side, the intermediary reference analyst usually has a better knowledge than the requester of the subject indexing codes and keywords (including how they have been used), and he can often translate the requester's needs into a more effectively worded query than if the requester is left to his own devices.

b. Direct Online Service

(1) If CIA's RECON data base is to be made available to all other NFIB agencies, there is a preferred alternative to merely expanding the operation described above. This would be to provide online access to the data base (stored at CIA Headquarters) via remote visual display terminals (VDTs) in other agencies. Such access could be made available on a 24-hour/day basis if necessary. Bibliographic references displayed on these remote VDTs could be printed immediately on medium-speed (300 lines/minute) printers co-located at each VDT. In this connection it should be pointed out that since the fall of 1973 a variety of intelligence analysts in CIA have been successfully querying the entire RECON data base directly via the SAFE Interim System¹ remote VDTs without OCR intervention. These analysts were formally trained to search the data base and are provided with guidance when necessary.

(2) The principal advantages of this arrangement include the significantly faster availability of the document citations to the analyst, plus the capability for the analyst to work directly with the data base. The latter feature would enable the analyst to determine if the subject codes and keywords he had chosen were producing references to the kinds of documents he needed; he could also see how large his document listing would be and modify his query parameters if necessary. All this could be done before ordering a printout from the system. For standing requests for index searches the capability to query the data base via the batch mode would be retained, rather than requiring the analyst to repeatedly compose his query at a terminal.

(3) If the online arrangement outlined is adopted, existing data communications systems such as the COINS network should be able to handle the transmission of the RECON bibliographic records from CIA Headquarters to requester terminals located at other NFIB agencies.

¹This is the precursor of the ultimate SAFE system, designed to assist in all aspects of intelligence production.

c. Online Service through Intermediaries

(1) Somewhere between options a. and b. above would be a system in which community customers would be linked to OCR's area reference analysts in a network of computer terminals. Queries would be presented telephonically or via the computer terminal, and the results of the analysts' online search could be displayed on the requester's terminal.

(2) The advantages of this blend of services are clear and have to do with effective, real-time communications between the area reference analyst and his customer. Questions about individual bibliographic references can be answered and the document listing tailored to the customer's needs. The refined listing could then be printed at the customer's printer as in option b.

4. Community Options for Document Retrieval Service:

a. Batch Mode

Under this configuration the CIA ADSTAR system would produce copies of documents after receiving requests either in writing or by computer terminal command, depending upon which form of bibliographic service has been adopted. The documents would be mailed to the requester.

b. Direct Online Retrieval

(1) In its most sophisticated configuration, remote ADSTAR terminals located throughout the Intelligence Community would allow non-CIA requesters to query the CIA's central ADSTAR library and display the text and print hard copies of whichever documents the NFIB analyst selected from his RECON listing.

(2) Such an online document retrieval system, however, could not be developed on the basis of existing data communications systems, such as the COINS network. This is because the bandwidth capacity to handle ADSTAR document image transmissions, which consists of approximately four million bytes per page image, is not available

Approved For Release 2002/01/08 : CIA-RDP83-00573R000100420011-9
transmission problem could be eased somewhat by using advanced data compression techniques, but even such a compressed data transmission would require an estimated one million bytes per page image.

5. Costs:

a. Any expansion of RECON services will require a major redesign of the data base. This redesign, to remove Input/Output bottlenecks and to render RECON capable of responding efficiently to larger online system requirements, would cost an estimated \$250,000, plus annual maintenance of \$100,000. These costs are basic and will be incurred if any major increase in the use of RECON is planned, whichever options are adopted.

b. If option 3.a. is adopted, about ten more document indexers and dissemination personnel would be needed to process the additional material expected to be added to the data base, in addition to indexing certain categories of documents in greater depth to satisfy the anticipated specific needs of various agencies. An additional typist would be necessary for the added input to the data base. Two additional camera operators would be needed in OCR's Microform Processing Branch to handle the increased volume of incoming documents to be filmed. Fifteen more area reference analysts would be needed to handle the added volume of requests.² At least two more clerks would be needed to address and package listings for mailing and to prepare document and courier receipts. Two additional direct access storage units (one primary and one backup) and one channel address unit would have to be purchased at a cost of \$175,000 in order to store the greater number of document citations in the data base. No additional computer equipment, software, personnel or floor space would be required. Operating expenses would probably approximate \$600,000 per year.

c. If option 3.b. is adopted (and existing communications systems are used), about half of the operating expenses cited in para. 5.b. above would be avoided, for the 15 area reference analysts would not be needed. A large, dedicated host computer would have to be installed, however, at a cost close to \$4 million. System software would have to be modified to make the computer program "reentrant," an arrangement enabling the central processing unit to handle

²It is extremely difficult to accurately estimate the number of index search requests that would be levied on CIA if RECON were made available to the Community without restriction. However, for the purpose of this memo, it is assumed that the current level of requests would increase five-fold. (This figure is largely a guess, based partly on OCR's experience with non-CIA use of RECON.)
Approved For Release 2002/01/08 : CIA-RDP83-00573R000100420011-9 their use of the RECON data base.)

up to 50 online requesters simultaneously. This would entail a one-time payment to a contractor, and would require approximately three man-years of his work and one calendar-year of time. An extra programmer and technician would each be needed in OCR's computer support unit to work with the contractor during the software modification and later to maintain this software and troubleshoot the system's operation. In addition to making the host computer operational for RECON, a number of other tasks would be required. The software interfaces connecting the computer, the message processor, and the COINS network would have to be developed. Certain additional software and hardware changes would be needed to adapt the RECON system to accommodate an increased number of users. Also, some combination of software modifications and human intervention may be required to resolve security release problems. Total cost for this effort would approximate \$500,000.

d. To house the host computer approximately 2,500 square feet of computer-grade floor space would be required, and ten positions would be needed for the personnel to operate the computer in a stand-alone environment that is electrically isolated from CIA's other computer facilities. The annual operating costs would include an additional computer programmer, and a computer technician, plus higher equipment maintenance costs. The total of these operating costs is estimated to be about \$220,000 per year for personnel and \$120,000 for maintenance.

e. In addition to the extra personnel--including indexers and microphotographers--already mentioned, a centralized staff of about three or four people (\$60-80,000/year) would probably be necessary to coordinate new indexing requirements from participating agencies; to train personnel to use the system and to provide on-going guidance once the system enters operation; and to handle trouble calls and transmit questions to appropriate operating personnel.

f. Option 3.c. would avoid the costs related to the installation and operation of a host computer and the attendant software development costs referred to in para. c. above, but the use of computer terminals to deliver bibliographic information would entail careful systems design and probably the acquisition of a number of "smart" terminals for use by OCR's analysts, terminals with the

ability to store information received from RECON and to deliver it on command to the remote customer terminal, which, in this configuration, would not have direct access to the CIA computer housing the RECON data above. Cost figures for such a system cannot be developed without a major study, but the costs should be significantly lower than those associated with the stand-alone host computer.

g. The costs of Document Retrieval Service Option 4.a. can also be separated into investment and operating expenses. An ADSTAR system augmented to provide Community-wide service would require approximately eight more storage modules to accommodate the assumed 25 percent increase in the number of documents five years old or less that are to be stored in that portion of the system designed to provide immediate retrieval. (These need not be added all at once; two per year could probably take care of the expected annual ADSTAR file growth.) Larger central processing units would be needed to accommodate the greater number of index records and associated support files. For the same reasons more disk packs and disk drives would be needed, the buffer capacity would have to be doubled and at least one other high-speed printer would have to be acquired. If this new centralized document service were to result in a demand for more documents in microfiche, the microfiche output capability would have to be greatly enhanced. Finally, software modifications to the ADSTAR system would be needed. These would all be one-time investment costs, and, while extremely conjectural, would probably total over \$1,000,000.

h. The increased operating costs anticipated for an expanded ADSTAR system would include two additional personnel to intervene in the ADSTAR process to resolve document release questions. Two extra clericals would be needed for packaging, mailing, and preparing document and courier receipts for batch requests for documents. Maintaining the various expanded support files (e.g., MIS and Security Access) would require another full-time employee. For preventive maintenance of the additional equipment, the maintenance contract would cost more. These operating costs would probably come to about \$150,000 per year.

i. Direct Online Retrieval, as in Option 4.b., would require additional outlays of \$750,000 for a central processing unit of greater capacity and associated support equipment, plus \$750,000 for more software, and (most importantly) the communications system hardware; the latter would include the communication lines themselves as well as the interface equipment, cryptographic systems, and remote access and display stations. Also, as with the online bibliographic retrieval system, appropriate measures would have to be taken to handle security release problems before this system is implemented. We cannot estimate the total of these additional costs without tasking communications specialists to undertake a system study, but undoubtedly the costs would be substantial.

j. It must be emphasized that the various costs described above are only preliminary estimates, subject to change. They are summarized in the tables attached to this memorandum.

6. Funding: There are no resources in the CIA Program for enhancement of our bibliographic index and document storage and retrieval capabilities beyond our immediate needs. If, after its study, the IHC validates a requirement to provide RECON and/or ADSTAR capabilities to other Community agencies and tasks CIA with the development, implementation, operation, and/or maintenance of these enhancements, then the IHC and the Resource Management Staff will have to identify the necessary resources. The resources required to expand and upgrade the existing system to serve the needs of other Community agencies should be provided by those agencies.

7. Time Required for Implementation: a. Any planned expansion of the CIA's bibliographic and document retrieval system would require a thorough and detailed study of at least six months' duration, plus time to hire whatever additional personnel the study will have called for.

b. Off-line bibliographic service (option 3.a.) could be implemented as soon as additional service personnel were hired, possibly as early as six months after completion of the initial six-month preliminary study, assuming that the requisite floor space could be acquired.

c. The more advanced approach of providing online bibliographic access (option 3.b.) would probably require at least two years after completion of the initial six-month study. During this period, software modifications would have to be accomplished, additional equipment would

have to be acquired and installed, and non-CIA agencies would have to program their budgets for the communications equipment and remote terminals they must fund. About the same time would be required to implement a system of online service through Intermediaries using a network of computer terminals (option 3.c.).

d. Centralized document retrieval would be impossible for the CIA until after the ADSTAR system had been implemented and operationally tested for at least six months. This would make ADSTAR available for Community-wide use no earlier than June 1980, and then only for batch retrieval (option 4.a.).

e. An online ADSTAR system that serviced non-CIA agencies via remote work stations (option 4.b.) would take at least two more years for programming user-agency budgets, and acquiring and installing the necessary additional equipment. FY 1982 would be a conservative target date.

8. Recommendation: a. We recommend that the IHC sponsor a study in depth of the Community's bibliographic and document retrieval needs to determine whether centralized services of the kinds described above would serve the Community's interests. The study should emphasize user requirements, system architecture (including communications), and precise investment and operating costs, together with offsetting savings to be made by reducing on-going activities or planned new ventures for which substantial expenditures are planned. Other aspects of the proposal which need research are the security restrictions to be imposed, and floor space requirements for machines and people.

b. If this study demonstrates that centralized services are desirable and economical, we recommend the adoption of RECON and ADSTAR in whichever of the configurations described above most effectively meets the needs of the Community, provided a suitable answer can be found to the questions of manning and funding the Community support.


Clifford D. May, Jr.

STATINTL

Att: a/s

PRELIMINARY ESTIMATES OF COSTS OF COMMUNITY BIBLIOGRAPHIC SYSTEM
(BASED ON RECON)

<u>Requirement</u>	<u>Option 3.a.</u>			<u>Option 3.b.</u>			<u>Option 3.c.</u>		
	<u>Positions</u>	<u>One-Time</u>	<u>Recurring</u>	<u>Positions</u>	<u>One-Time</u>	<u>Recurring</u>	<u>Positions</u>	<u>One-Time</u>	<u>Recurring</u>
Redesign RECON		250,000	100,000		250,000	100,000		250,000	100,000
<u>Bibliographic Service</u>									
<u>Off-line</u>									
- 13 Index/Dissem/Clerical, 2 Camera Op., 15 Area Reference Analysts	30		600,000	15		300,000	30		600,000
- Add. Direct Access Storage Unit		175,000			175,000			175,000	
<u>On-line (Direct)</u>									
- Host Computer					3,200,000*				
- Software					500,000				
- 10 Operators, 1 Tech, 1 Systems Analyst, 3 Requirements Coord.				15		280,000			
- Operating Costs						120,000			
<u>On-line (Intermediary)</u>									
- Smart Terminals								250,000	
- Software								250,000	
Sub-Totals	30	425,000	700,000	30	4,125,000*	800,000	30	925,000	700,000
Total Annual Cost Assuming 5-Year System Life			85,000** \$785,000			825,000** \$1,625,000			185,000** \$885,000

*Plus 2500 sq. ft. of floor space.
**Annual figures represent 1/5 of the one-time totals shown in preceding column.

PRELIMINARY ESTIMATES OF COSTS OF COMMUNITY DOCUMENT RETRIEVAL SYSTEM

<u>Requirement</u>	<u>Option 4.a.</u>			<u>Option 4.b.</u>		
	<u>Positions</u>	<u>One-Time</u>	<u>Recurring</u>	<u>Positions</u>	<u>One-Time</u>	<u>Recurring</u>
Hardware (storage modules, CPU, disk drives, buffer, printer and software)		1,000,000			1,000,000	
Maintenance			150,000			150,000
Document Release Control	2		40,000	2		40,000
Clerical Service	2		25,000			
Files Support	1		20,000	1		20,000
Additional ADSTAR Hardware, Software					1,500,000	100,000
Communications				Unknown	Unknown	Unknown
Sub-Totals	5	1,000,000	235,000	3	2,500,000	310,000
Total Annual Cost Assuming 5-Year System Life		<div> <div>→</div> <div>200,000*</div> <div>\$435,000</div> </div>			<div> <div>→</div> <div>500,000*</div> <div>\$810,000</div> </div>	

*Annual figures represent 1/5 of the one-time totals shown in preceding column.

22 OCT 1979

MEMORANDUM FOR: Chairman IHC and Members of the
IHC Bibliographic Study Group

SUBJECT: Preliminary Draft Report

1. As the contractor for the IHC Bibliographic Study, we have formulated some conclusions/recommendations based upon our analysis of the CIA proposal, the information gathered to date during the course of the study, and our discussion with you. These are:

Conclusions:

- o General reductions in the resources required for manpower support and computing functions related to bibliographic services and document storage and retrieval could be achieved if this system were implemented. This should be the first step in a series of Community-wide endeavors to improve the flow of automated intelligence support within the Community.

Recommendations:

- o Accept CIA's proposal to provide limited direct on-line access to the RECON bibliographic data base by intelligence analysts throughout the Community via COINS.
- o Give further study to the proposal to expand the scope of the current RECON system to cover gaps in existing or proposed bibliographic systems.
- o Allow a limited number of users with proper clearances within each agency of the Community to have total access to the full RECON bibliographic data base. These should primarily be information specialists in the central reference services of the various user agencies. Access to the full text of the document can and should be limited on the basis of need-to-know.
- o Develop a plan for a total distributed Community bibliographic system in which the RECON system would constitute a principal component.

Derivative C1 By 064718
Review 15 Oct 1999
Reason: CGB3

SECRET/NOFORN

Approved For Release 2002/01/08 : CIA-RDP83T00573R000100120011-9

- o Give further study to the ways and means by which the CIA's ADSTAR system can contribute to a total Community bibliographic system.
- o Take action to ensure that the bibliographic index system being developed as a part of the SAFE project eventually becomes an integral part of the Community bibliographic system.

2. The attached document is supplied to you to summarize some of the results of the study and our principal recommendations as to the conclusions which should be drawn therefrom. It is designed to form the basis for further discussion by the study group and the formulation of a report of the group itself to the IHC at the conclusion of this phase of the study. The final report should also address areas requiring follow-on study, particularly in relation to the feasibility of the ADSTAR portion of the CIA proposal.

3. We look forward to receiving your suggestions and comments and to assisting you in formulating a report from the study group to the IHC.

25X1A



Attachment: a/s

Approved For Release 2002/01/08 : CIA-RDP83T00573R000100120011-9

SECRET/NOFORN

Approved For Release 2002/01/08 : CIA-RDP83T00573R000100120011-9

PRELIMINARY CONCLUSIONS
FOR THE STUDY OF
COMMUNITY BIBLIOGRAPHIC SERVICES (U)

15 October 1979

Derivative C1 By 064718
Review 15 Oct 1999
Reason: CGB3

Approved For Release 2002/01/08 : CIA-RDP83T00573R000100120011-9

COMMUNITY BIBLIOGRAPHIC SERVICES

BACKGROUND

1. (S) The June 1979 report of the Analyst Support Task Force which was sponsored by the DCI's Intelligence Information Handling Community (IHC) concluded that "a community-wide, on-line, multi-source bibliographic index to Intelligence Community documents should be created and made available either directly or through an information systems specialist" to the intelligence community's production analysts. The task force report also indicated that varying intra-agency distribution policies and procedures prevent analysts in one agency from having access to the same source materials that their colleagues in the other agencies see and use. Analysts feel they are not getting access to all intelligence source materials required to provide assessments in assigned areas of responsibility. Such practices impede the coordination process and may result in uninformed analysis.

2. (U) In response to the analyst support study, CIA has submitted a proposal to the IHC to make its all-source bibliographic reference file (RECON) available to the community as a centralized intelligence community reference system that would be centrally managed and operated for the community by CIA. The RECON system currently provides on-line access to bibliographic references from all intelligence sources and on all subjects of intelligence interest to CIA analysts in the Langley headquarters building, NPIC, and the NITC. The classification of the materials referenced ranges from UNCLASSIFIED to TOP SECRET SI/TK/G.

3. (U) In addition to its RECON bibliographic service, CIA offered its automated document storage and retrieval (ADSTAR) system as a centralized community document storage and retrieval system. ADSTAR is part of CIA's SAFE project. CIA proposed three options for the community use of the RECON bibliographic reference system and two options for the ADSTAR document storage and retrieval service which together over a 5 year period would range in cost from \$6.05 million to \$12.18 million. ^{1/}

^{1/} The top end figure did not include costs for secure wide-band communications or microfilm. Approved For Release 2002/01/08 : CIA-RDP83T00573R000100120011-9

An additional 33 to 35 positions would be required to support this effort under CIA's proposal, but the job functions to be performed by these people would vary depending on the options chosen. CIA's options for bibliographic services were: (1) Increased off-line service to the community with requests mailed to CIA for action, (2) Direct on-line service to the RECON computer by non-CIA NFIP analysts, and (3) on-line service through a CIA intermediary. CIA proposed services for automatic document storage and retrieval were: (1) off-line service requested by mail, or (2) direct on-line service to CIA's ADSTAR system by non-CIA NFIP analysts. 2/

CURRENT STUDY

4. (S) An IHC-sponsored Bibliographic Study Group is nearing completion of its analysis of CIA's proposed services. This group studied the feasibility of the CIA proposal by comparing the capabilities of 6 of the 30 identified bibliographic systems currently maintained in the community. The study included an analysis of a bibliographic system at the Department of State. The group's initial findings are as follows:

- A direct comparison of sources maintained by the various agencies in automated form indicates that RECON contains more references to more intelligence sources than does any other automated bibliographic system evaluated in this study; however, it contains references to only 50% of all intelligence sources maintained by the DIA, NSA, NPIC, and State Department. The most notable shortfalls were in U.S. SIGINT field reports, State Cables, and Open Source Literature.
- RECON's automated index record and coding scheme is similar to that used by DIA in its three major bibliographic files, but NPIC's EPF, NSA's SOLIS, and the AF CIRC II systems have all moved towards a version of full text retrieval of entire documents (disseminated electrically) or abstracts of larger documents. Such full text systems provide a high degree of analyst's direct interaction through the use of an English-like retrieval language. Currently, many of the community bibliographic systems (including RECON) make the user translate his requests into abbreviated numerical codes primarily designed for the computer.
- RECON maintains references on more restricted intelligence source documents (namely ORCCN and

2/ As noted earlier, this option did not include costs for secure wide-band communications of microfilm stored source documents.

GAMMA) than any other system; however, most of the other agencies maintain some of this material in their automated systems. Under CIA's RECON proposal, many of these restricted materials would not be made available to non-CIA NFIP users. In 1979 these restricted materials amounted to approximately 23% of the RECON data base and this figure was growing at the rate of approximately 6% per year based on the study team's trend analysis for the past 2-1/2 years.

- o The RECON updating process is much like that used by DIA, State, and NPIC (i.e., very labor intensive). The NSA SOLIS and Air Force CIRC II systems have partially automated the indexing and data reduction processes required to maintain their bibliographic systems. The codes used and the structure of the data bases at DIA are similar to RECON.
- o Direct on-line access to the RECON data base is currently restricted to CIA analysts in the Langley Headquarters building, NPIC, and the NITC. As a base of comparison, the locations of all other terminals served by the evaluated bibliographic system at DIA, NSA, NPIC, State, and the Air Force were counted to determine the number of additional terminals a centralized RECON would provide. Over 375 terminals would have to be added to the system unless it was linked through the COINS network. Using the COINS network (as proposed by CIA under option b), less than 100 terminals would need to be added in order to provide equivalent access to RECON. If direct access via COINS were initially limited to users in the central reference groups of each agency, less than 20 additional terminals would have to be added.
- o While direct on-line access to the RECON data base is currently limited to CIA analysts, several agencies currently make use of the data base and have expressed interest in obtaining direct on-line access to it. For the past several years CIA management has placed severe limits on other agency access to the RECON system because of substantial reductions imposed on CIA resources. However, the bibliographic references to finished intelligence products from RECON have been transported to NSA and loaded on the NSA COINS host

SECRET/NOFORN

computer to provide partial RECON access by the community's production analysts. This subset of the RECON data base available through COINS was queried by non-CIA NFIP analysts over 1,200 times in 1978 and over 1,900 times during the first six months of 1979. Over 1,600 off-line requests for searches to the full RECON data base were made by non-CIA NFIP agencies in 1978. In 1974 the assistant Secretary of Defense for Intelligence (ASD/I) sponsored a study of the Community On-Line Intelligence System (COINS) which recommended that CIA's RECON/AEGIS bibliographic index be made accessible throughout the community via the COINS network. This recommendation was never implemented. The non-CIA NFIB use of RECON has been growing over the years. This indicates the continued community interest in accessing the RECON data base.

- o CIA's proposed option for indirect electronic bibliographic services through an intermediary (CIA option 3c) was deemed infeasible since nearly \$1 million would have to be spent for hardware procurement or software services and an additional 15 area specialists would be required to receive and handle requests from analysts outside CIA. Technically this option would be very risk because the capability to provide such a service does not currently exist. In addition, this proposed option did not call for an electronically isolated processor. Apparently the smart terminals identified by CIA in its proposal would provide the electronic isolation CIA requires under option 3b. It seems doubtful that state-of-the-art in smart terminals has advanced to the degree that they could provide such a service.

5. (U) Congress has repeatedly questioned the overall direction and planning of ADP in the intelligence community. From reviews of the SAFE requirements documentations and during discussions with the joint CIA/DIA SAFE project managers concerning the role of bibliographic support (namely RECON) in the SAFE project, it was evident that CIA SAFE plans to include an enhanced version of RECON in mid-1983; however, current SAFE milestones do not call for the enhanced RECON SAFE to be shared either with DIA or the rest of the intelligence community. The project managers were aware that CIA had proposed to make RECON available to the community. Both project managers agreed that if the CIA proposal were adopted numerous technical, political, administrative, and policy

problems that could hamper progress on the SAFE project would be solved or at least initially addressed, particularly if the CIA SAFE bibliographic index should be made accessible outside CIA. They looked upon the effort as an investment which would provide benefits to the SAFE project.

PROBLEMS

6. (U) CIA's current RECON proposal is a means for providing a larger segment of the intelligence community with common access to a broad section of intelligence community source documents. The RECON system by itself will not provide references to 100% of the source documents, but it could provide an interim solution to what has been analyzed as a significant problem within the community. The major problem is that the intelligence community lacks a long-range plan for its bibliographic services. Parts of the problem are being addressed through major projects such as SAFE, SOLIS, WEEDER, PROJECTOR, and COINS, but very little overall coordination and direction have been applied to this effort.

7. (C) Distributed Bibliographic Responsibilities. Current RECON processes and procedures are too labor-intensive to be used as a totally centralized bibliographic reference and document storage and retrieval system for the community. Based on the bibliographic study group's analysis, 60 to 75 additional people would have to be added to the CIA staff to provide centralized community services. While much of the manpower for such an effort could be drawn from other agencies (assuming that their bibliographic systems would be eliminated), this does not seem to be a feasible alternative for either political or technical reasons. Technically, other bibliographic systems currently used in the community are much more widely available, use English-like text to provide more analyst interaction, and are less labor-intensive to maintain. The disadvantage to these systems is that they require much more of an investment in hardware and software. Such an investment would likely be in the neighborhood of \$20 million and require at least 25 additional personnel. The most feasible solution to the bibliographic problem would appear to be one of delegated responsibility for data base maintenance and standardized practices across the community. Such an approach envisions that intelligence source document originators would not only be responsible for the development of source materials, but also for the timely and accurate maintenance of their assigned portions of the "community's bibliographic data bank." Under such an approach, NSA's SOLIS system could provide community-wide references to SIGINT materials, the Air Force CIRC II system could provide references to open source S&T information, and RECON/SAFE could provide references to both CIA and DoD IIRs and finished intelligence products. State cables could be

provided either through NSA's project WEEDER or through CIA/DIA SAFE. The most feasible alternative for providing community-wide access to such systems would appear to be the COINS network.

8. (S) Assuming that this is the basis for a long range plan to provide community bibliographic services, several security and policy issues must be addressed. The bibliographic study group noted that CIA's current RECON has more GAMMA and ORCON material than does any other system (the NSA SOLIS system contains no GAMMA material). Under its proposal, CIA feels that it can isolate the GAMMA and ORCON materials into files within the RECON data base and restrict access to such material through the use of software and hardware access controls. Those CIA management personnel that developed the proposal quickly note that computer security experts at CIA would have to grant final approval for such an approach. Based on the following facts, it would seem that a direct on-line access to the RECON data base could be achieved with a minimum security risk:

- CIA insists that an electronically isolated processor would have to be procured to process the community-wide accessible RECON data base. Such a requirement would seem appropriate since the current RECON processor in the RUFFING computer complex is electronically attached to three other mainframes that process highly sensitive technical collection materials. In addition, CIA has approached the security issue much like NSA in that it feels protection can be maintained with isolated host processors attached through "secure" communications processors. (The CIA RUFFING center is similar in capability and configuration to [REDACTED])

25X1A

- Initially, only a limited number of users would be granted access to the RECON data base (the study group estimates 20 initial users). These users should be part of the current central reference groups in each agency. They should be cleared to the TS/SI/TK/G level and it would appear appropriate that these people should be polygraphed.
- The structure of the RECON data base provides its own level of protection since it displays only a bibliographic reference to a source document.

SECRET/NOFORN

Approved For Release 2002/01/08 : CIA-RDP83T00573R000100120011-9

The paucity of information in RECON (particularly English-like text) has been sighted as a major disadvantage for interactive use by intelligence analysts; however, this community-wide "limited" access to the data base can be viewed as a significant security advantage. It would allow increased community access with a minimal risk. Other systems in the community which provide English-like text of entire documents or abstracts (i.e., NSA, SOLIS and AF CIRC II) do not currently provide on-line access to highly sensitive documents (or their bibliographic references) because the envisioned security risk is too great.

9. (S) Assuming that the policy and security issues favor the use of RECON as an interim approach for the eventual development of a "distributed" community bibliographic system, certain administrative problems must still be resolved. First, CIA has indicated that it has no resources to provide community bibliographic services. CIA estimated that it would need over \$10 million to provide on-line bibliographic services and off-line document storage and retrieval support for a community system. The requirement for a new electronically isolated processor and required software would cost \$4.125 million according to the CIA proposal. The study group analysis has determined that a new processor, some additional required software, and initial COINS linkage could be achieved for under \$3.5 million. According to the CIA proposal the balance of the \$10 million cost estimate (5.875 million) would be required for additional automated document storage and retrieval equipment and for personnel expenses over a 5 year period. Approximately 35 additional personnel would be required under the CIA proposal. The study group concluded that under this interim solution CIA should need at most \$215,000 for additional ADSTAR equipment and no more than 5 additional people for ADSTAR operation over the next 5 years. The total 5 year cost for a new processor, COINS communications, additional ADSTAR equipment, and 15 personnel would be just over \$5 million (assuming 1979 dollars). In principle, that figure could be lowered by \$1.7 million for the new processor and 4 personnel (\$400,000 over five years) because the study group analysis indicates that only 15-50% of the present processor is used by RECON. The excess capacity in the new machine should be used to enhance RECON, provide additional community automatic services, or be integrated into the SAFE project. It has been suggested that this excess capacity could be used to assist in the processing of Soviet Defector and emigre reports.

10. (U) The second major administrative problem is space for the new processor. CIA says it has no additional floor space for the new system. CIA estimated that 2,500 square

Approved For Release 2002/01/08 : CIA-RDP83T00573R000100120011-9
feet would be required. The size of the floor space required was apparently based on RECON's current IBM 370/168 configuration. More powerful, compact, and less expensive computer equipment can be procured for the RECON system thereby reducing this space problem (i.e., an AMDAHL 470 V/6 needs only 33% of the floor space that IBM 370/168 requires, and the AMDAHL needs no plumbing for cooling purposes).

CONCLUSIONS

11. (U) General reductions in the resources required for manpower support and computing functions related to bibliographic services and document storage and retrieval could be achieved if this system were implemented. This should be the first step in a series of community-wide endeavors to improve the flow of automated intelligence support within the community.

RECOMMENDATIONS

12. (U) The basic CIA proposal for centralized bibliographic services would cost about \$5 million over the next 5 years (in 1979 dollars), but it would provide an interim solution to a significant problem noted in the recent survey of intelligence production analysts. It would also highlight technical and policy problems (and provide initial solutions) that will hamper other major automated community projects now being developed. The following recommendations are made concerning the CIA proposal:

- o Accept CIA's proposal to provide limited direct on-line access to the RECON bibliographic data base by intelligence analysts throughout the community via COINS.
- o Give further study to the proposal to expand the scope of the current RECON system to cover gaps in existing or proposed bibliographic systems.
- o Allow a limited number of users with proper clearances within each agency of the community to have total access to the full RECON bibliographic data base. These should primarily be information specialists in the central reference services of the various user agencies. Access to the full text of the document can and should be limited on the basis of need-to-know.

UNCLASSIFIED

UNCLASSIFIED

Approved For Release 2002/01/08 : CIA-RDP83T00573R000100120011-9

- o Develop a plan for a total distributed community bibliographic system in which the RECON system would constitute a principal component.
- o Give further study to the ways and means by which the CIA's ADSTAR system can contribute to a total community bibliographic system.
- o Take action to ensure that the bibliographic index system being developed as part of the SAFE project eventually becomes an integral part of the community bibliographic system.

Approved For Release 2002/01/08 : CIA-RDP83T00573R000100120011-9

UNCLASSIFIED

~~SECRET~~ NOFORN

Approved For Release 2002/01/08 : CIA-RDP83T00573R000100120011-9

Description of Funding Alternatives
and Manpower Requirements

- 1(a) Provides for the acquisition of an Amdahl V-5 CPU or equivalent together with required peripherals, on a five-year lease purchase, outright purchase of COINS network access system hardware, software development costs and operation and maintenance funds for commencement of operations in the second half of FY 81, as follows:

	<u>81</u>	<u>82</u>	<u>83</u>	<u>84</u>	<u>85</u>
o V-5 CPU--5 yr lease purchase	624	624	624	624	624
o Peripherals--5 yr lease purchase	186	186	186	186	186
o COINS NAS	270	0	0	0	0
o Software	575	100	100	100	100
o Operations & Maintenance	<u>100</u>	<u>400</u>	<u>400</u>	<u>400</u>	<u>400</u>
	1,760	1,310	1,310	1,310	1,310

- 1(b) Provides for the return to the original plans for replacing the CAMS 370/158 back-up CPU with an Amdahl V-6 processor on a five-year lease purchase beginning in FY 81 instead of FY 82 and utilizing the 370/158 to support RECON. All other funding would remain the same as in 1(a) except that a network access system funded in the COINS FY 80 budget would be diverted from other possible uses to support of RECON.

	<u>81</u>	<u>82</u>	<u>83</u>	<u>84</u>	<u>85</u>
o Peripherals--5 yr lease purchase	186	186	186	186	186
o Software	575	100	100	100	100
o Operations & Maintenance	<u>100</u>	<u>400</u>	<u>400</u>	<u>400</u>	<u>400</u>
Total RECON	861	686	686	686	686
CAMS CPU	<u>750</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>
Total Additions to NFIP	1,611	686	686	686	686

Approved For Release 2002/01/08 : CIA-RDP83T00573R000100120011-9

~~SECRET~~ NOFORN

SECRET NOFORN

Approved For Release 2002/01/08 : CIA-RDP83T00573R000100120011-9

- 1(c) Provides for minimum software development in FY 81 utilizing \$100K in FY 80 RMS contract funds to support effort if necessary, utilizes FY 80 CAMS NAS and defers purchase of RECON equipment to FY 82 on four-year lease purchase.

	<u>81</u>	<u>82</u>	<u>83</u>	<u>84</u>	<u>85</u>
o Software	375	100	100	100	100
o Operations & Maintenance	0	400	400	400	400
o V-5 CPU--4 yr lease purchase	0	708	708	708	708
o Peripherals--4 yr lease purchase	<u>0</u>	<u>220</u>	<u>220</u>	<u>220</u>	<u>220</u>
	375	1,428	1,428	1,428	1,428

- 1(d) Provides minimum software development with RMS contract support as in 1(c), utilizes surplus 370/158 CAMS back-up CPU and COINS FY 80 NAS, with peripherals purchased on a four-year lease purchase.

	<u>81</u>	<u>82</u>	<u>83</u>	<u>84</u>	<u>85</u>
o Software	375	100	100	100	100
o Operations & Maintenance	0	400	400	400	400
o Peripherals--4 yr lease purchase	<u>0</u>	<u>220</u>	<u>220</u>	<u>220</u>	<u>220</u>
	375	720	720	720	720

It should be noted that cost projections and manpower requirements are projected without allowance for inflation or for support of new or additional demands which will likely be made upon the system because of its increased availability.

In Options 1(a) and 1(b), it is assumed that equipment procurement and software completion, installation, test and integration will require approximately six months. Manpower requirements are stated in manyears, but the total number of positions required will approach 15 during the second half of fiscal 1981.

SECRET NOFORN
Approved For Release 2002/01/08 : CIA-RDP83T00573R000100120011-9

SECRET

Approved For Release 2002/01/08 : CIA-RDP83T00573R000100120011-9

Manpower estimates are based upon 24 hour a day operation requiring 10 operators, at least two for systems programming and programming maintenance and three additional employers in OCR to handle document printing, wrapping and mailing in response to expected increases in demand for these reviews generated by the increased use of RECON. This latter requirement has not been projected over time and will, in all probability, grow if provisions for electrical distribution of documents from the ADSTAR system are not made. If the proposed system is not also utilized to support CIA and if COINS manpower is not increased to permit 24 hour operation, manpower requirements for operations personnel will be significantly lower.

SECRET

Approved For Release 2002/01/08 : CIA-RDP83T00573R000100120011-9

ROUTING AND RECORD SHEET

SUBJECT: (Optional)

STATINTL

FROM:

DDA/Management Staff
7D18 HQS

EXTENSION

NO.

DATE

31 OCT 1979

STATINTL

TO: (Officer designation, room number, and building)

DATE

RECEIVED

FORWARDED

OFFICER'S
INITIALS

COMMENTS (Number each comment to show from whom to whom. Draw a line across column after each comment.)

1. Director of Data
Processing

TUBE DX-6

11/3

J

2.

3.

DD

5 NOV

CJ

4.

C/M

5.

6.

Cma -
File in P.B.G.

7.

(X-ref in IHC file)

8.

9.

10.

11.

12.

13.

14.

15.

Per our discussion,
attached is the final RMS
issue paper on the Agency's
proposal to IHC on an
expanded RECON.

STAT

Att: a/s

no action by
you is required
T